

Gauri Toshniwal

E-mail: gaurigst1970@gmail.com | LinkedIn: [GauriToshniwal](#) | Github: [GauriToshniwal](#) | Mob: 9130687967

Objective: Seeking Opportunities in Machine Learning (NLP)

EDUCATION

National Institute of Technology, Tiruchirappalli Tamil Nadu, India
Masters of Technology - Data Analytics (CGPA: 9.36/10) 08/2023 - 06/2025
Coursework: Deep Learning, NLP, Machine Learning, Federated Learning, Mathematics (Linear Algebra, Probability)

Government College of Engineering, Nagpur Maharashtra, India
Bachelor of Engineering - Computer Science and Engineering (CGPA: 9.42/10) 06/2021
Coursework: Operating Systems, Data Structures & Algorithms, Artificial Intelligence, Databases, Discrete Mathematics, Compilers

SKILLS

Languages and Frameworks: Python, C++, C; Pandas, Scikit-learn, PyTorch, Numpy, Flask, Pyspark, Spacy, Pyredis
Tools and Platforms: Git, HuggingFace, SQL, MongoDB, ElasticSearch, Jupyter Notebooks, Google Colab

EXPERIENCE

Findability Sciences Remote, India
Data Scientist (Onsite - Hindustan Times Digital) 10/2021 - 08/2023

- **Demographics (Gender, Age, Interest) Prediction:** User Level Classifiers - Developed multiple classifiers using grammatical and entity based features from news headlines with average model **f1-score of 0.65**. Deployed these models in production to serve **2 million** users daily.
- **Subscription Churn Prediction and Retention Feed:** Built an XGB classifier model to predict subscriber churn. Utilized SHAP to explain the model's features and provide insights into the factors that were driving subscriber churn. Developed retention feed based on subscription drivers identified.
- **User Stitching - Data Engineering task:** Developed a function to replace existing user history API to reduce the load on the server due to API hits, Implemented logic for identifying and stitching user history for a unique user visiting from multiple devices/Platforms.
- **Premium Story Feed [Recommend Story with high likelihood to subscribe]:** Created a premium stories feed by analyzing the user journey from non-subscriber to the subscriber. Achieved **5x** improvement in **CTR** through the implementation of the premium feed.

Data Analyst Remote, India
05/2021 - 10/2021

- **Project Timeline and Human Resource Forecasting (Clustering - Regression) :**
Predicting Total Efforts in Man Hours and Total Duration of the project in months required by different types of resources.
- **Operation Theatre Occupancy Duration Prediction (Regression):**
Developed the prediction model using ridge regression to support operation theater scheduling and optimization.

PROJECTS

- **Math Word Problem Solving** 07/2024 - Current
Pursuing a thesis on innovative approaches to math word problems.
- **Closed Domain Event Extraction on RAMS Dataset** [Github](#) 05/2024
Transformed event trigger identification, classification and argument detection, into token classification problems using BIO tagging, then applied BERT-based models for token classification.
- **Tree Based Federated Learning (PoC)** [Blog](#), [Github](#) 05/2024
Basic Cross-Silo Horizontal Federated Learning PoC using Socket Programming with Decision Tree-based Model Training
- **Comprehensive Solution for Indian Sign Language Recognition and Translation:** [Github](#) 06/2021
Developed an ISL recognition and translation system, creating video and hand landmark datasets, and implementing an English-to-ISL text module using PoS tagging and rule based rearranging. Utilized Inception v3 and LSTM for converting ISL gestures to English text, and Mediapipe for real-time ISL detection and classification.

ACADEMIC ACHIEVEMENTS

- **Awarded Reliance Foundation Postgraduate Scholarship for the year 2023-24** 06/2024
Received by 100 top students nationwide for exceptional academic achievement and potential in technology and engineering
- **ACM Summer School on Responsible and Safe AI** [Certificate](#) 06/2024
Selected and Attended summer school hosted by IIT Madras
- **AI Safety Fundamentals - Bluedot Impact** [Certificate](#) 03/2024 - 06/2024
Selected for a 12-week online course, covering a range of technical AI alignment research agendas
- **DECOR 2022@ICDE, 38th IEEE International Conference on Data Engineering** [Paper](#) 05/2022
Won Best Student Paper Award for "Object Detection in Indian Food Platters using Transfer Learning with YOLOv4"